Vector Autoregressions

• VAR: Vector AutoRegression
  – Nothing to do with VaR: Value at Risk (finance)
• Multivariate autoregression
• Multiple equation model for joint determination of two or more variables
• One of the most commonly used models for applied macroeconometric analysis and forecasting in central banks
Two-Variable VAR

- Two variables: $y$ and $x$
- Example: output and interest rate
- Two-equation model for the two variables
- One-Step ahead model
- One equation for each variable
- Each equation is an autoregression plus distributed lag, with $p$ lags of each variable
VAR(p) in 2 Variables

\[ y_t = \mu_1 + \alpha_{11} y_{t-1} + \alpha_{12} y_{t-2} + \cdots + \alpha_{1p} y_{t-p} + \beta_{11} x_{t-1} + \beta_{12} x_{t-1} + \cdots + \beta_{1p} x_{t-p} + e_{1t} \]

\[ x_t = \mu_2 + \alpha_{21} y_{t-1} + \alpha_{22} y_{t-2} + \cdots + \alpha_{2p} y_{t-p} + \beta_{21} x_{t-1} + \beta_{22} x_{t-1} + \cdots + \beta_{2p} x_{t-p} + e_{2t} \]
Multiple Equation System

• In general: $k$ variables
• An equation for each variable
• Each equation includes $p$ lags of $y$ and $p$ lags of $x$
• (In principle, the equations could have different # of lags, and different # of lags of each variable, but this is most common specification.)
• There is one error per equation.
  – The errors are (typically) correlated.
Unrestricted VAR

- An unrestricted VAR includes all variables in each equation
- A restricted VAR might include some variables in one equation, other variables in another equation
- Old-fashioned macroeconomic models (so-called simultaneous equations models of the 1950s, 1960s, 1970s) were essentially restricted VARs
  - The restrictions and specifications were derived from simplistic macro theory, e.g. Keynesian consumption functions, investment equations, etc.
VAR Revolution

• Christopher Sims
  – Princeton University
  – 2011 Nobel Prize in Economics
• “Macroeconomics and Reality” (1980)
  – Sims argued that conventional macro models were “incredible” – they were based on non-credible identifying assumptions
Sims and VARs

• Sims argued that the conventional models were restricted VARs, and the restrictions had no substantive justification
  – Based on incomplete and/or non-rigorous theory, or intuition
• Sims argued that economists should instead use unrestricted models, e.g. VARs
• He proposed a set of tools for use and evaluation of VARs in practice.
Estimation

- Each equation estimated by OLS

\[
y_t = \mu_1 + \alpha_{11}y_{t-1} + \alpha_{12}y_{t-2} + \cdots + \alpha_{1p}y_{t-p} \\
+ \beta_{11}x_{t-1} + \beta_{12}x_{t-1} + \cdots + \beta_{1p}x_{t-p} + e_{1t}
\]

\[
x_t = \mu_2 + \alpha_{21}y_{t-1} + \alpha_{22}y_{t-2} + \cdots + \alpha_{2p}y_{t-p} \\
+ \beta_{21}x_{t-1} + \beta_{22}x_{t-1} + \cdots + \beta_{2p}x_{t-p} + e_{2t}
\]
Estimation in Stata

• To estimate a VAR in the variables \( y \) & \( x \) with lags 1 through \( p \) included
  – .varbasic \( y \) \( x \), lags(1/\( p \))

• For example, using readgdpgrowth.dta and variables \( gdp \) and \( d.t1year \) with 3 lags
  – .gen rate=d.t1year
  – .varbasic rate \( gdp \), lags(1/3)

• Could also use
  – .var rate \( gdp \), lags(1/3)
Example: GDP and Interest Rate

```
. varbasic rate gdp, lags(1/3)

Vector autoregression

Sample: 1954q2 - 2014q4
No. of obs = 243
Log likelihood = -862.264
AIC = 7.212049
FPE = 4.647489
HQIC = 7.293109
Det(Sigma_ml) = 4.141545
SBIC = 7.413296

Equation Parms RMSE R-sq chi2 P>chi2
rate 7 0.651329 0.2131 65.81013 0.0000

No. of obs = 243
Log likelihood = -862.264
AIC = 7.212049
FPE = 4.647489
HQIC = 7.293109
Det(Sigma_ml) = 4.141545
SBIC = 7.413296

Equation Parms RMSE R-sq chi2 P>chi2
rate 7 0.651329 0.2131 65.81013 0.0000
"
## Interest Rate Equation

|     | Coef.   | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-----|---------|-----------|-------|-------|---------------------|
| rate|         |           |       |       |                     |
| L1. | .2368376| .0636494  | 3.72  | 0.000 | .1120871            | .361588            |
| L2. | -.3440265| .063117  | -5.45 | 0.000 | -.4677335           | -.2203195          |
| L3. | .2837125| .0655516  | 4.33  | 0.000 | .1552337            | .4121914           |
| gdp |         |           |       |       |                     |
| L1. | .043465 | .0129585  | 3.35  | 0.001 | .0180667            | .0688632           |
| L2. | .0244206| .0131164  | 1.86  | 0.063 | -.0012871           | .0501282           |
| L3. | -.0132412| .0129672 | -1.02 | 0.307 | -.0386564           | .0121739           |
| _cons| -.1759707| .0680549 | -2.59 | 0.010 | -.3093559           | -.0425856          |
## GDP Equation

|   | Coef.  | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|---|--------|-----------|------|------|----------------------|
| gdp |        |           |      |      |                      |
| rate |        |           |      |      |                      |
| L1. | .1921948 | .3247687 | 0.59 | 0.554 | -0.4443401 to 0.8287297 |
| L2. | -1.391911 | .3220522 | -4.32 | 0.000 | -2.023122 to -0.7607008 |
| L3. | .0099461 | .3344749 | 0.03 | 0.976 | -0.6456127 to 0.6655059 |
| gdp |        |           |      |      |                      |
| L1. | .2949096 | .0661204 | 4.46 | 0.000 | 0.165316 to 0.4245033 |
| L2. | .1948721 | .0669259 | 2.91 | 0.004 | 0.0636998 to 0.3260444 |
| L3. | -.0096068 | .0661645 | -0.15 | 0.885 | -0.1392868 to 0.1200733 |
| _cons | 1.657327 | .3472477 | 4.77 | 0.000 | 0.9767339 to 2.33792   |
Order Selection

• A VAR(p) includes $p$ lags of each variable in each equation
• In a two-variable system, the number of coefficients in each equation is $1+2p$
  – The total number is $2(1+2p)=2+4p$
• In a $k$-variable system, the number of coefficients in each equation is $1+kp$
  – The total number is $k(1+2p)=k+2kp$
• How should $p$ be selected?
• Common approach:
  – Information criterion, primarily AIC
AIC and BIC for VAR Models

\[ AIC = -2L + 2(k + 2kp) \]

\[ BIC = -2L + (k + 2kp)\ln(T) \]

where \( L \) is log-likelihood from model

• Select model with smallest AIC (or BIC)
Stata Implementation

- **varsoc command**
- To calculate information criterion for a VAR in variables x and y up to a maximum lag of $p_{max}$:
  - `.varsoc x y, maxlag(pmax)`
- Produces a convenient table
Example: GDP and Interest Rate

`. varsoc rate gdp, maxlag(8)`

Selection-order criteria
Sample: 1955q3 - 2014q4  
Number of obs = 238

<table>
<thead>
<tr>
<th>lag</th>
<th>LL</th>
<th>LR</th>
<th>df</th>
<th>p</th>
<th>FPE</th>
<th>AIC</th>
<th>HQIC</th>
<th>SBIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-889.718</td>
<td></td>
<td></td>
<td></td>
<td>6.15753</td>
<td>7.49343</td>
<td>7.50519</td>
<td>7.52261</td>
</tr>
<tr>
<td>1</td>
<td>-871.849</td>
<td>35.739</td>
<td>4</td>
<td>0.000</td>
<td>5.48013</td>
<td>7.37688</td>
<td>7.41216</td>
<td>7.46442</td>
</tr>
<tr>
<td>2</td>
<td>-855.383</td>
<td>32.931</td>
<td>4</td>
<td>0.000</td>
<td>4.93516</td>
<td>7.27213</td>
<td>7.33093</td>
<td>7.41802*</td>
</tr>
<tr>
<td>3</td>
<td>-845.741</td>
<td>19.285*</td>
<td>4</td>
<td>0.001</td>
<td>4.70672*</td>
<td>7.22471*</td>
<td>7.30703*</td>
<td>7.42896</td>
</tr>
<tr>
<td>4</td>
<td>-843.448</td>
<td>4.5846</td>
<td>4</td>
<td>0.333</td>
<td>4.77493</td>
<td>7.23906</td>
<td>7.3449</td>
<td>7.50167</td>
</tr>
<tr>
<td>5</td>
<td>-839.439</td>
<td>8.0184</td>
<td>4</td>
<td>0.091</td>
<td>4.77485</td>
<td>7.23898</td>
<td>7.36834</td>
<td>7.55995</td>
</tr>
<tr>
<td>6</td>
<td>-838.662</td>
<td>1.5543</td>
<td>4</td>
<td>0.817</td>
<td>4.90635</td>
<td>7.26607</td>
<td>7.41894</td>
<td>7.64539</td>
</tr>
<tr>
<td>7</td>
<td>-833.965</td>
<td>9.3935</td>
<td>4</td>
<td>0.052</td>
<td>4.87828</td>
<td>7.26021</td>
<td>7.43661</td>
<td>7.69789</td>
</tr>
<tr>
<td>8</td>
<td>-832.04</td>
<td>3.8495</td>
<td>4</td>
<td>0.427</td>
<td>4.96486</td>
<td>7.27765</td>
<td>7.47756</td>
<td>7.77369</td>
</tr>
</tbody>
</table>
Result

- For this example
  - AIC selects $p=3$
  - BIC selects $p=2$

- Notice that the AIC value for $p=3$ in this table (AIC=7.224) is different from that obtained when we estimated the VAR(3) model (AIC=7.212).
  - This is because for the AIC comparison, all estimates are from a common sample, in this case excluding the first 8 observations since the maximum order is set to 8

- The varsoc command is correct
Double Check

. varbasic rate gdp if time>=tq(1955q3), lags(1/3)

Vector autoregression

Sample: 1955q3 - 2014q4
No. of obs = 238
Log likelihood = -845.7406
AIC = 7.224711
FPE = 4.706718
HQIC = 7.307028
Det(Sigma_ml) = 4.184176
SBIC = 7.428962

• When we constrain the sample to exclude the first 8 observations, the reported AIC is 7.224, correctly.
Let’s look at the VAR(3) estimates again.
### Example: GDP and Interest Rate

|      | Coef.   | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|------|---------|-----------|-------|-------|----------------------|
| rate | rate    |           |       |       |                      |
| L1.  | 0.2368376 | 0.0636494 | 3.72  | 0.000 | 0.1120871, 0.361588  |
| L2.  | -0.3440265 | 0.063117  | -5.45 | 0.000 | -0.4677335, -0.2203195 |
| L3.  | 0.2837125 | 0.0655156 | 4.33  | 0.000 | 0.1552337, 0.4121914 |
| gdp  |         |           |       |       |                      |
| L1.  | 0.043465 | 0.0129585 | 3.35  | 0.001 | 0.0180667, 0.0688632 |
| L2.  | 0.0244206 | 0.0131164 | 1.86  | 0.063 | -0.0012871, 0.0501282 |
| L3.  | -0.0132412 | 0.0129672 | -1.02 | 0.307 | -0.0386564, 0.0121739 |
| _cons| -0.1759707 | 0.0680549 | -2.59 | 0.010 | -0.3093559, -0.0425856 |
| gdp  |         |           |       |       |                      |
| L1.  | 0.1921948 | 0.3247687 | 0.59  | 0.554 | -0.4443401, 0.8287297 |
| L2.  | -1.391911 | 0.3220522 | -4.32 | 0.000 | -2.023122, -0.7607008 |
| L3.  | 0.0099461 | 0.3344749 | 0.03  | 0.976 | -0.6456127, 0.665505  |
| gdp  |         |           |       |       |                      |
| L1.  | 0.2949096 | 0.0661204 | 4.46  | 0.000 | 0.165316, 0.4245033  |
| L2.  | 0.1948721 | 0.0669259 | 2.91  | 0.004 | 0.0636998, 0.3260444  |
| L3.  | -0.0096068 | 0.0661645 | -0.15 | 0.885 | -0.1392868, 0.1200733 |
| _cons| 1.657327  | 0.3472477 | 4.77  | 0.000 | 0.9767339, 2.33792    |
Interpretation

• It is difficult to interpret the large number of coefficients in the VAR model

• Main tools for interpretation
  – Impulse responses
Impulse Response Analysis

- VAR(1) with no intercept
  \[ y_t = \alpha_{11} y_{t-1} + \beta_{11} x_{t-1} + e_{1t} \]
  \[ x_t = \alpha_{21} y_{t-1} + \beta_{21} x_{t-1} + e_{2t} \]

- The impulse responses are the time-paths of to y and x in response to shocks.
Impulse Response Analysis

• The errors may be correlated.
• We “orthogonalize” them

\[ e_{1t} = u_{1t} \]
\[ e_{2t} = \rho e_{1t} + u_{2t} \]
\[ = \rho u_{1t} + u_{2t} \]
Orthogonalized Model

\[ y_t = \alpha_{11} y_{t-1} + \beta_{11} x_{t-1} + u_{1t} \]

\[ x_t = \alpha_{21} y_{t-1} + \beta_{21} x_{t-1} + \rho u_{1t} + u_{2t} \]

- The shocks \( u_1 \) and \( u_2 \) are uncorrelated
- The ordering matters
  - The shock to \( y \) affects both \( y \) and \( x \) in period \( t \)
  - The shock to \( x \) affects only \( x \) in period \( t \)
- The impulse responses are the time-paths of \( y \) and \( x \) in response to the shocks \( u_1 \) and \( u_2 \)
- Imagine \( y=0 \) and \( x=0 \). Set \( u_1 = 1 \). Trace the history of \( y \) and \( x \)
Impulse Responses by Recursion

\[ y_1 = \alpha_{11}0 + \beta_{11}0 + 1 = 1 \]

\[ x_1 = \alpha_{21}0 + \beta_{21}0 + \rho 1 = \rho \]

\[ y_2 = \alpha_{11}y_1 + \beta_{11}x_1 = \alpha_{11} + \beta_{11} \]

\[ x_2 = \alpha_{21}y_1 + \beta_{21}x_1 = \alpha_{21} + \beta_{21} \rho \]

\[ y_3 = \alpha_{11}y_2 + \beta_{11}x_2 = \alpha_{11}(\alpha_{11} + \beta_{11}) + \beta_{11}(\alpha_{21} + \beta_{21} \rho) \]

\[ x_3 = \alpha_{21}y_2 + \beta_{21}x_2 = \alpha_{21}(\alpha_{11} + \beta_{11}) + \beta_{21}(\alpha_{21} + \beta_{21} \rho) \]
Impulse Responses

- The impulse responses are these time-paths of $y$ and $x$ due to the shocks $u_1$ and $u_2$
- They are found by this recursion formula
- They are functions of the estimated VAR coefficients
Impact of Shocks on Variables

• In a 2-variable system, there are 4 impulse response functions
  – The effect on y of a shock to y (u₁)
  – The effect on y of a shock to x (u₂)
  – The effect on x of a shock to y (u₁)
  – The effect on x of a shock to x (u₂)

• In a k-variable system, there are $k^2$ impulse response functions!
Stata Calculation

• Impulse response automatically calculated with varbasic command
• A kxk matrix of impulse response is created
GDP/Interest Rate Example

Graphs by irfname, impulse variable, and response variable
Interpretation

• Labeled “Graphs by irfname, impulse variable, and response variable”
  – “Impulse variable” means the source of the shock
  – “Response variable” means the variable being affected
• Upper left: “varbasic, gdp, gdp”
  – Impact of a gdp shock on the time-path of gdp
• Upper right: “varbasic, gdp, rate”
  – Impact of a gdp shock on the time-path of interest rates
• Lower left: “varbasic, rate, gdp,”
  – Impact of an interest rate shock on the time-path of gdp
• Lower right: “varbasic, rate, rate”
  – Impact of an interest rate shock on the time-path of interest rates
• The impulse response is graphed as a function of forward time periods
Scale

• The graphs are all created on the same scale, so difficult to read
• It may be better to create graphs separate for each impulse response
  . irf graph oirf, impulse(gdp) response(rate)
• This creates the impulse response for the impact of a gdp shock on the time-path of interest rates
GDP on GDP

```
.irf graph oirf, impulse(gdp) response(gdp)
```

Graphs by irfname, impulse variable, and response variable
GDP on Interest Rates

. irf graph oirf, impulse(gdp) response(rate)
Interest Rates on Interest Rates

```
.irf graph oirf, impulse(rate) response(rate)
```

Graphs by irfname, impulse variable, and response variable.
Interest Rates on GDP

`. irf graph oirf, impulse(rate) response(gdp)`
3-variable system

- Interest Rate Change (12-month T-Bill)
- Investment Growth Rate
- GDP Growth Rate
GDP/Investment/Interest Rate

Graphs by irfname, impulse variable, and response variable
Investment Shock on GDP

Graphs by irfname, impulse variable, and response variable
Investment Shock on Interest Rate

Graphs by irfname, impulse variable, and response variable
Interest Rate Shock on Investment

Graphs by irfname, impulse variable, and response variable